

Review of the 2006 Summer Flounder Assessment Update

Individual Reviewer's Report

Prepared for the NOAA National Marine Fisheries Service

By

Dr. Joseph E. Powers
Louisiana State University
2147 Energy Coast and Environment Building
Baton Rouge, LA 70803

At the request of the Science and Research Director of the NMFS Northeast Fisheries Science Center and the NMFS Office of Science and Technology, I was asked to participate in a review of the summer flounder stock assessment update.

The Terms of Reference for the review were to:

- 1) Review the draft document "Summer Flounder Biological Reference Point Update for 2006" by Mark Terceiro (BRP2006_5.doc, dated Aug. 10).
- 2) Provide comments and recommendations regarding:
 - a) Is an appropriate historical time period being used to provide biological inputs for the projections?
 - b) Has an appropriate adjustment been made for the assessment model's documented retrospective bias?
 - c) Is the rebuilding target and rate based upon an accurate estimate of the recruitment levels expected as the stock rebuilds?
- 3) Provide comments and recommendations regarding possible future improvements in the assessment of summer flounder.

In addition to myself, two NMFS scientists (exclusive of Northeast Fisheries Science Center (NEFSC) scientists) participated in the review: Dr. Richard Methot, NMFS Office of Science & Technology, Review Panel (Chair of the discussions) and Dr. Owen Hamel, NMFS Northwest Fisheries Science Center. The discussions took place at the Northeast Fisheries Science Center, Woods Hole, MA; Sept 14-15, 2006

The report documented below represents my findings and recommendations from that review.

1. Review the draft document "Summer Flounder Biological Reference Point Update for 2006" by Mark Terceiro (BRP2006_5.doc, dated Aug. 10).

The above reference point update, the assessment document (NEFSC Ref Doc 06-17) and associated documents and files were provided to me more than a week prior to the meeting. Additionally, the methods and results were presented to the Review Panel by Dr. Mark Terceiro at the meeting.

The assessment has been conducted using ADAPT vpa model fitting routines using in excess of 40 indices of abundance and estimates of catch-at-age from age-length keys. The history of the fishery and its management results are best exemplified by the Figure 14 in NEFSC 06-17, i.e. the "control rule" figure. Since implementation of management measures in the late 1990's the assessment indicates that fishing mortality was reduced and the population responded with an increase in stock biomass; but in the last few years fishing mortality appears to have stabilized and, thus, the improvement in stock biomass has stalled. These trends alone would indicate that recent catches have probably not been sufficiently reduced to keep the stock on a rebuilding trajectory within specified time frames.

While not taking away from the basic results of the assessment, there are a number of issues (discussed in Term of Reference 2, below) which impact the rebuilding target and how it is measured, catches needed to rebuild by the year 2010 and the fishing rate at the target.

2. Provide comments and recommendations regarding:

- a. Is an appropriate historical time period being used to provide biological inputs for the projections?
- b. Has an appropriate adjustment been made for the assessment model's documented retrospective bias?
- c. Is the rebuilding target and rate based upon an accurate estimate of the recruitment levels expected as the stock rebuilds?

a. Is an appropriate historical time period being used to provide biological inputs for the projections?

Rebuilding criteria (Minimum Stock Size Threshold or MSST) and the target stock size are defined in the Technical Guidelines. Implicit in that definition is the notion that the ideal metric for "stock size" is in units of reproductive potential. Often we in the stock assessment profession are lax in our terminology in reference to stock size: sometimes we use total biomass, sometimes spawning biomass and sometimes egg production. What we are trying to measure is reproductive potential; thus, sometimes total biomass is the best measure we have available for reproductive potential, sometimes spawning biomass and sometimes egg production. In the case of summer flounder, evidence has accumulated to indicate the need to move from a total biomass metric to a spawning biomass metric. This change allows the use of maturity data that indicates a rather rapid maturity schedule.

Additionally, the spawning weight to be used should be more closely aligned within the year with the actual season in which spawning occurs. Thus, a January 1 time for calculating weight should not be used for age zero. Note, that this change does not affect the VPA fitting and basic stock trends. It only affects the interpretation of those trends in terms of some relatively small changes in spawning biomass.

NEFSC 06-17 and the associated files indicated that weight at age has changed in recent years. Therefore, we discussed the appropriateness of using the recent weights-at-age versus the earlier weights-at-age when conducting projections. Recall that there are two types of projections which are required: the first are short term projections in order to evaluate the status of the stock up to the year 2010 under various management actions; the second set of projections are longer term evaluating management relative to the stocks long term productivity potential, i.e. MSY. Therefore, the former set of projections is best made assuming that biological characteristics like weight-at-age are similar to the present values, i.e. using the 2003-2005 body weights. Whereas, long term projections should reflect the range of biological outcomes that may occur, i.e. the longer 1982-2005 data set.

b. Has an appropriate adjustment been made for the assessment model's documented retrospective bias?

The assessment and subsequent reviews have noted a retrospective pattern in the assessments in which subsequent assessments show higher fishing mortality rates for the same years and ages than earlier assessments AND redoing the assessment in which the terminal years are lopped off one by one shows the same pattern. These retrospective patterns are common in assessments and are due to the mismatch of time series data to the model being fitted. Thus, they are expected in a statistical sense; yet, they have important implications in the assessment. However, a large number of factors can cause these patterns (change in catchability of indices, misreporting of catch, etc). Therefore, it is difficult to make an adjustment without knowing what it is that needs fixing. Additionally, in the most recent year, the retrospective pattern in summer flounder appears to have diminished to be inconsequential. Also, earlier in the summer flounder time series another retrospective pattern existed but in the opposite direction. It too disappeared after a few years. We did some testing by redoing the assessment with higher catches than were reported in recent years and (subsequent to the meeting) by substituting a small number for the observed zeros in the indices instead of treating zeros as missing data. The model fits reacted predictably, lowering the effect of the retrospective pattern. I concur that the assessment should include not treat observed zeros as missing data and the solution suggested for substituting a small value is appropriate at this stage. However, this aspect should be treated more rigorously in the future.

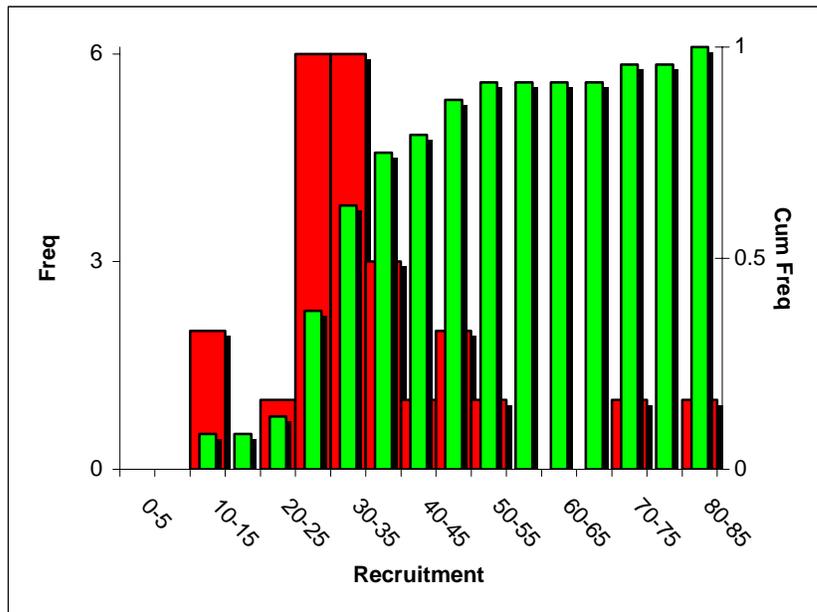
Other than that recommendation, there does not appear to be an appropriate adjustment available for the summer flounder retrospective pattern. Therefore, I suggest none. However, managers should be cognizant of this issue when setting management measures.

c. Is the rebuilding target and rate based upon an accurate estimate of the recruitment levels expected as the stock rebuilds?

This question is essentially asking: what are the productivity characteristics of the stock when it is recovered and what will they likely be over the next few years until 2010?

The time series of recruitment from the assessment showed the first five or so years of recruitment being variable but at a high level followed by about 20 years of being variable at a lower level. The 2005 estimate was low (but not the lowest on record) and within the normal variability exhibited in this time series.

Although stock-recruitment models were fit to the data the results were less than satisfactory. Therefore, empirical approaches have been suggested to which I concur. The empirical approach is to assume that future recruitment will be similar to the pattern that exists in the time series data. The question then is: when the stock biomass recovers will it be producing recruitment like in the early 1980's or more like it is producing now. In order to address both of these issues I recommend, as did the rest of the Panel) that a stochastic method of choosing future recruitments be used, rather than just using the mean or median with an assumed distribution. The stochastic method randomly picks values from the cumulative frequency distribution of the observed recruitments. See the figure:



Note that these data were taken from the original assessment, so the values change a little with the “zero” adjustment, but the argument is still the same. The figure indicates that the distribution is slightly skewed with there being more likelihood that a lower value will occur in a single year.

This stochastic method coupled with these data has two desirable properties: for short term projections (i.e. until 2010) it is more likely that three years of recruitment will be low on the scale so the consequences to the population is more likely that recruitment

will be similar to current levels. Using the same stochastic selection methodology, the consequences to stock biomass of long term projections will be more like average recruitment. Thus, the stochastic methodology is somewhat self weighting and not inconsistent with short and long term perceptions of productivity.

This same line of reasoning suggests that the recruitment level used to calculate the benchmark biomass at F_{max} should be the average recruitment rather than the median.

3. Provide comments and recommendations regarding possible future improvements in the assessment of summer flounder.

- a) the ADAPT methodology is fairly structured in the case of summer flounder: 41 indices and the ADAPT assumption that catch at age is known with certainty do not leave much flexibility in fitting. Therefore, it is hard to explore other structures to try to get at changing q 's, changing partial recruitment, changing M 's etc. Therefore, it is recommended that the various forward projecting models that have become available be explored. In particular, these models usually allow some flexibility in the choice of the probability model being used in a particular maximum likelihood component. This may be especially useful in using the zero data in indices. The newer models may be able to incorporate survey data using distributions that allow zeros.
- b) Weighting of indices is a complicated and detailed debate which has not been resolved satisfactorily in many arenas (for example ICCAT). However, the use of 41 indices as if they were independent may be stretching things. Statistical methods to appropriately group indices should be explored.
- c) The reproductive biology of summer flounder appears to be both important to the assessment and somewhat unique (38% of age 0's spawn, most of the younger fish samples are male dominated while most of the older samples are female dominated). These have implications for natural mortality rates and productivity in general. Existing samples should be analyzed more fully and the feasibility of improving collection in the future should be evaluated.
- d) Methods to improve catch estimates should be developed. This may mean enforcement methods need to be developed, as much as scientific methods. Discard estimates should also be improved. Improvements in this area would help to alleviate retrospective problems.